SEQUENCE LISTING

<110> Fritig, Bernard
      Toquin, Valerie
      Geoffroy, Pierrette
      Legrand, Michel
      Kauffmann, Serge

<120> INDUCIBLE COMTII PROMOTER, CHIMERA GENE
   CONTAINING SAME AND TRANSFORMED PLANTS

<130> A34638-PCT-USA-I (072667.0189)

<140> To Be Assigned
<141> 2003-08-04

<150> US 09/937,204
<151> 2001-12-13

<150> PCT/FR00/00714
<151> 2000-03-22

<150> France 99/03700
<151> 1999-03-22

<150> France 99/07646
<151> 1999-06-11

<160> 26

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 1863
<212> DNA
<213> Nicotiana tabacum

<220>
<221> misc_signal
<222> (667)...(672)
<223> Inverted W box

<221> misc_RNA
<222> (820)...(830)
<223> Inverted L box

<221> enhancer
<222> (845)...(852)

<221> misc_signal
<222> (1034)...(1047)
<223> P box

<221> misc_signal

<222> (1221)...(1226)
<223> G box

<221> misc_signal
<222> (1343)...(1356)
<223> Inverse L box

<221> misc_signal
<222> (1369)...(1374)
<223> A box

<221> misc_signal
<222> (1377)...(1382)
<223> GT box

<221> misc_signal
<222> (1483)...(1488)
<223> GT box

<221> misc_signal
<222> (1562)...(1567)
<223> Inverted W box

<221> misc_signal
<222> (1600)...(1614)
<223> L box

<221> CAAT_signal
<222> (1675)...(1679)

<221> misc_signal
<222> (1681)...(1690)
<223> E box

<221> CAAT_signal
<222> (1695)...(1699)

<221> TATA_signal
<222> (1735)...(1739)

<221> misc_signal
<222> (1772)...(0)
<223> Transcription origin

<400> 1
aaagttaggg acaatctata gtgtcacaaa gttgcttatg gcttttggtt cagataaaga 60
aaaagaacag catttttaatt tgtgaagatt agtctgagca gaatttcatt gtatctagaa 120
agaaattgaa aaaagaaata ttctatttca ctattatgtt aggtgcaact atatcatcac 180
catggaaaag ccggagtaaa aagagaacgt agaggagatt tcatgatttg attgagaata 240
taatatatta ttttttttgta attccacaca aagattaaga aaatgatctg atcaatgatg 300
gctccgagga tttggctgtc gcgggaacta tgacattaat ataaatttgt cgctgcctat 360
aaagacccta tctatctatc tatctatcta tatatatata tatatatata tatatatata 420
tatatatata tatatatata tatatatata tatatataag cgctaatatt tgattatttt 480
ttaaaaatat ttataagtat atatgaaatt tttgacgaaa tttttgtgtg accgtgaccc 540
ctcaacctat agtgtgcgtc cacctgtgcc aacaatatag agacaatttg ctcgtatagt 600
cagaaagagt gttttacttt ttagttgctt tttagtgaat ctactcggta taaagttaaa 660
ttagtgggtc aataagtcgg gtgaatagtt aaagaaaaca gtggtgagtt tagctgtcaa 720

```
ataatttctt cttttcttg ttttcacatt agaaatcaaa ataaaacaca agctttttgt 780
atttatttta acacaagcta attatatgtt tatatgctgg ttaggtgaag taaagcatgt 840
tatatgagga aagtacgaag aaaatgtgcc aattgtcgtg tacagcaaag cagccagcac 900
aagcaaattc gcacttgata agtggctaag tccactttct agtggaccta gtggttcact 960
aactttacc aaaaaggcaa taatttgcaa ttcaaaaaga aaaaaggaaa aaagaaaact 1020
agacagactt taacacacca actcccacag gaagcaacaa tgcaactcac aaaaggaaac 1080
cgagttttc cgcgacggat ctagaatttg ggttcattct ttacgctttt tcgtattaaa 1140
ctcattatat ttgtataatt atgggtttat attttttatt tattgtaatt tttgtaaaat 1200
tttatatata agtgtatact ccacgtctcc ggatactaca ttagcctcta gggttcttaa 1260
tactcttgtt aaattgtcca ggctccaaac gcatgttcgt ttcaatttta acggatgttt 1320
ccgaacaact ccaaatgttc aatgttaggt gtgtttggtg ttaagcttcc gtcctaggtt 1380
aatagaatag ataattgttg tttcttatat agttttgaac aatcgtcgcc ataaactaat 1440
ttttaggatg gaagctaatt tttaggatgg agtacagcct aaggttaaaa tataactata 1500
aaaaatatcc ataaaaggtg aaatttaatt agtaacatga aaagataaaa ctagtgttat 1560
cggtcaaact ttcaaaagag aaagaaataa ctagacaaac ttcaacaacc aacctgccca 1620
acatgctact gtgcaattga aaaataaaca aaagagaacc agacaatatt tcaaccaata 1680
ttccatcaag aaaaccaatt atgacaattc ttaaccaaag tcacaactaa cacttataaa 1740
aagcactaac tcaactgtac atgattgtga agcctaacaa aaacactcta aaaggaaaag 1800
actacgagaa taattacact acaactctta tagctaattc ttgtctcaag attttcagct 1860
atg                                                          1863

<210> 2
<211> 5371
<212> DNA
<213> Nicotiana tabacum

<220>
<221> promoter
<222> (1)...(1860)

<221> misc_signal
<222> (1772)...(0)
<223> Transcription origin

<221> intron
<222> (2282)...(3633)

<221> intron
<222> (3945)...(4726)

<221> terminator
<222> (5090)...(5371)

<400> 2
aaagttaggg acaatctata gtgtcacaaa gttgcttatg cttttggtt cagataaaga 60
aaaagaacag cattttaatt tgtgaagatt agtctgagca gaatttcatt gtatctagaa 120
agaaattgaa aaaagaaata ttctatttca ctattatgtt aggtgcaact atatcatcac 180
catggaaaag ccggagtaaa aagagaacgt agaggagatt tcatgatttg attgagaata 240
taatatatta tttttttgta attccacaca aagattaaga aaatgatctg atcaatgatg 300
gctccgagga tttggctgtc gcgggaacta tgacattaat ataaatttgt cgctgcctat 360
aaagacccta tctatctatc tatctatcta tatatatata tatatata tatatatata 420
tatatatata tatatatata tatatatata tatataag cgctaatatt tgattatttt 480
ttaaaaatat ttataagtat atatgaaatt tttgacgaaa tttttgtgtg accgtgaccc 540
ctcaacctat agtgtgcgtc cacctgtgcc aacaatatag agacaatttg ctcgtatagt 600
cagaaagagt gttttacttt ttagttgctt tttagtgaat ctactcggta taaagttaaa 660
ttagtgggtc aataagtcgg gtgaatagtt aaagaaaaca gtggtgagtt tagctgtcaa 720
ataatttctt cttttcttg ttttcacatt agaaatcaaa ataaaacaca agctttttgt 780
```

```
atttatttta acacaagcta attatatgtt tatatgctgg ttaggtgaag taaagcatgt 840
tatatgagga aagtacgaag aaaatgtgcc aattgtcgtg tacagcaaag cagccagcac 900
aagcaaattc gcacttgata agtggctaag tccactttct agtggaccta gtggttcact 960
aacttttacc aaaaaggcaa taatttgcaa ttcaaaaaga aaaaggaaa aaagaaaact 1020
agacagactt taacacacca actcccacag gaagcaacaa tgcaactcac aaaaggaaac 1080
cgagtttttc cgcgacggat ctagaatttg ggttcattct ttacgctttt tcgtattaaa 1140
ctcattatat ttgtataatt atgggtttat attttttatt tattgtaatt tttgtaaaat 1200
tttatatata agtgtatact ccacgtctcc ggatactaca ttagcctcta gggttcttaa 1260
tactcttgtt aaattgtcca ggctccaaac gcatgttcgt ttcaatttta acggatgttt 1320
ccgaacaact ccaaatgttc aatgttaggt gtgtttggtg ttaagcttcc gtcctaggtt 1380
aatagaatag ataattgttg tttcttatat agttttgaac aatcgtcgcc ataaactaat 1440
ttttaggatg gaagctaatt tttaggatgg agtacagcct aaggttaaaa tataactata 1500
aaaaatatcc ataaaaggtg aaatttaatt agtaacatga aaagataaaa ctagtgttat 1560
cggtcaaact ttcaaaagag aaagaaataa ctagacaaac ttcaacaacc aacctgccca 1620
acatgctact gtgcaattga aaaataaaca aaagagaacc agacaatatt tcaaccaata 1680
ttccatcaag aaaaccaatt atgacaattc ttaaccaaag tcacaactaa cacttataaa 1740
aagcactaac tcaactgtac atgattgtga agcctaacaa aaacactcta aaaggaaaag 1800
actacgagaa taattacact acaactctta tagctaattc ttgtctcaag attttcagct 1860
atggaatcct caaccaaaag ccaaatacca acacaatcag aagaagagcg taactgcaca 1920
tatgccatgc aactattgtc atcttcagtc ctcccctttg tgttgcattc aacaattcaa 1980
ttggaagttt ttgagatatt agccaaatct aatgacacta aactttctgc ttctcaaatt 2040
gtttctcaaa ttcctaactg cacaaaacct gaagcaccta ctatgttaaa taggatgctt 2100
tatgtcttgg ctagttactc cttgtttact tgttccattg ttgaagatga aaaaaataat 2160
gggggccaaa aaagagtgta tggtttgtca caagtgggaa aattctttgt taaaaatgaa 2220
aatggtgcat caatgggggcc actttggct ttgcttcaaa ataaagtatt cataaacagc 2280
tggtaagttt tgtcctactg tgtattcttt ttgcagtggc tgtattgatt ggttgccttt 2340
ttcacaagac aagattctta agtttatta cttgtcgatt tatgttagtc gtatgtgcta 2400
gtgttattat tctccatctg atcctttat tggtcacttt acctaaaaat attgttacaa 2460
aacatttgtc cttctagaaa atcaggtatt attaatttt caattccatc tttattactc 2520
caatagtgaa tatggttatt aattagtgtt ttaaggaaga tgtaaggata atttaatcaa 2580
ataggattta ttattaatgt tgtcaaagat tctggtggat ggatcggaga aaatttcttc 2640
atcttaatca gagtttgatg ttcgagccac aggaatgaat ttgtttttaa tagggagtat 2700
tttctctttg aatagacctt acacaataaa aggacaaccc ggtacactaa gcttccgtta 2760
tgcgcggggt tcggggaaag gaccgcatca ccaggtctat tgtacgcagc gttacccaac 2820
gtgaatctaa attaatgaga ctaaaaaatg gaacccaaca ccagtgaaaa ccaaaaaaag 2880
aagcaaactt tagtggatgg cttggaaaga tctttcttct tgaataactt ggagcgctat 2940
atattaaggc gtcgcagccg ttagatactt tcaagaagaa agctaaaaaa tgttttaaag 3000
ttacggcgct agaataatga aatttctcta tatatataat tcaaaagtta ataatttatt 3060
ctcttaactt aaatctatat tataaaacta tattaagtaa cttctgccta atttataata 3120
tacaactaat gttttgagaa aacaaaataa caacaacatc aaacccaatg aaatcccaca 3180
agtagagttt ggggaggata gtgtgtacgg agaccttacc cctaccttat aaagttaaag 3240
aggctgtttt cgaaagactc tcggctcaag aacattaaaa atttgagaaa acaaaatata 3300
aattcaaaac ctatattaag tttataatcc atggtatatt atattggctt agtaatctga 3360
aatgaaagat ttatgtttga ctcctctaaa cttgtttta atgcaaaaga ggcacaacat 3420
atatattata agtatctttt tttggtttcc cactgtggcc gctaaattcg gattcgctgg 3480
aagtgtcaca ttgttggaga tgggggcaac gctcacaaca aagacgattc tataattagt 3540
gttcgaacct gaaattttag ttaaagataa agaagtactt accataatgg tagatatgat 3600
catatctgac tctctttcta atttcaaatt acaggtttga actaaaagat gcagttcttg 3660
aaggaggagt tccatttgac agggtacacg gtgtgcatgc atttgaatat ccaaaatcgg 3720
acccaaaatt caatgatgtt ttcaacaagg caatgatcaa tcacacaact gtagtcatga 3780
aaaaaatact tgaaaattac aaaggttttg agaaccttaa aactttggtt gatgttggag 3840
gtggtcttgg agttaacctc aagatgatta catctaaata ccccacaatt aagggcacta 3900
attttgattt gccacatgtt gttcaacatg ccccttccta tcctggtacc ttctctcgtt 3960
cttattttgt tgtttattat atttacttcg atcatcaggt ctaggtctgt caagttaaat 4020
tcgttctcaa aaaagtttat aaaggttttg aactccatca cctattgctt taggattttg 4080
agttgtatgc tctgagtctt gcgcatggta tcatagtcaa tttatttaag ctcgttattg 4140
cacttgtgaa ttctattata taaggagtaa gcctaccaaa aaggagcgaa aatattttcc 4200
```

```
aaaactcttt ttaaaccttc ctcacccccat tccctctcc cctctcccc aacaccaccc 4260
accacccaa ctcccccgtc ttagttttt tatttatcct ggactttctt atattttatg 4320
ctttccttta attgaactct tgtaactaaa ccatttgccc cccaccctat agtgtttgcc 4380
taaatttat atttttcaaa ataatatttt ctatttacta attaaacatt agaaaatatt 4440
tttcggattt ttttccactc accaaccaag catgggaaaa tagtgataaa actactcatt 4500
tttcaaaata atattttcaa ggaaaacatt ttcctttata ccaaatacc ttactcttgt 4560
atacaaatct tcatgtcgat gatcttgcaa tatatataca tgtatatgta tgatttgata 4620
aaccacatga acaaaatggt tgagctctgc gaattgtgat atatgatttg cttatgtgtt 4680
gtgcactatc aattacttaa attaaacttc atctaataat attgcagggg tggaacatgt 4740
tggggagat atgtttgaaa gtgttccaga aggagatgct attttatga agtggattct 4800
tcatgactgg agtgatagtc acaacctcaa gttgctaaag aactgctaca aggctctacc 4860
agacaatgga aaggtgattg ttgttgaggc cattttacca gtgaaaccag acattgacac 4920
cgcagtggtt ggcgtttcgc aatgtgattt gatcatgatg gctcaaaatc ctggaggcaa 4980
agagcgatcg gaagaggagt ttcgagcctt ggctactgaa gctggattca aaggcgttaa 5040
cttaatatgt tgtgtctgta atttttgggt catggaattc tgcaagtaga tttctactgt 5100
acattgagtt tctactactc ttgagtatcc atttatggca atctgggact ggaattgcag 5160
cttagtccag attgaacatt gatattccta ataatatttc tattatttcc cttgtttatt 5220
tctcttgtat gaaaggatgt cattttgagt attgataatc atgttctcta ggacagaaat 5280
tgtaactttg tccaacttta ttgatattcc tagtaagatt tatatgacat gtgtctctgg 5340
tttgagaaga gtttcaatat ctacagacgg g                    5371
<210> 3
<211> 1095
<212> DNA
<213> Nicotiana tabacum

<220>
<221> CDS
<222> (1)...(1095)

<400> 3
atg gaa tcc tca acc aaa agc caa ata cca aca caa tca gaa gaa gag      48
Met Glu Ser Ser Thr Lys Ser Gln Ile Pro Thr Gln Ser Glu Glu Glu
 1               5                  10                  15

cgt aac tgc aca tat gcc atg caa cta ttg tca tct tca gtc ctc ccc      96
Arg Asn Cys Thr Tyr Ala Met Gln Leu Leu Ser Ser Ser Val Leu Pro
                20                  25                  30

ttt gtg ttg cat tca aca att caa ttg gaa gtt ttt gag ata tta gcc     144
Phe Val Leu His Ser Thr Ile Gln Leu Glu Val Phe Glu Ile Leu Ala
            35                  40                  45

aaa tct aat gac act aaa ctt tct gct tct caa att gtt tct caa att     192
Lys Ser Asn Asp Thr Lys Leu Ser Ala Ser Gln Ile Val Ser Gln Ile
        50                  55                  60

cct aac tgc aca aaa cct gaa gca cct act atg tta aat agg atg ctt     240
Pro Asn Cys Thr Lys Pro Glu Ala Pro Thr Met Leu Asn Arg Met Leu
 65                  70                  75                  80

tat gtc ttg gct agt tac tcc ttg ttt act tgt tcc att gtt gaa gat     288
Tyr Val Leu Ala Ser Tyr Ser Leu Phe Thr Cys Ser Ile Val Glu Asp
                85                  90                  95

gaa aaa aat aat ggg ggc caa aaa aga gtg tat ggt ttg tca caa gtg     336
Glu Lys Asn Asn Gly Gly Gln Lys Arg Val Tyr Gly Leu Ser Gln Val
                100                 105                 110
```

```
gga aaa ttc ttt gtt aaa aat gaa aat ggt gca tca atg ggg cca ctt        384
Gly Lys Phe Phe Val Lys Asn Glu Asn Gly Ala Ser Met Gly Pro Leu
        115                 120                 125

ttg gct ttg ctt caa aat aaa gta ttc ata aac agc tgg ttt gaa cta        432
Leu Ala Leu Leu Gln Asn Lys Val Phe Ile Asn Ser Trp Phe Glu Leu
    130                 135                 140

aaa gat gca gtt ctt gaa gga gga gtt cca ttt gac agg gta cac ggt        480
Lys Asp Ala Val Leu Glu Gly Gly Val Pro Phe Asp Arg Val His Gly
145                 150                 155                 160

gtg cat gca ttt gaa tat cca aaa tcg gac cca aaa ttc aat gat gtt        528
Val His Ala Phe Glu Tyr Pro Lys Ser Asp Pro Lys Phe Asn Asp Val
                165                 170                 175

ttc aac aag gca atg atc aat cac aca act gta gtc atg aaa aaa ata        576
Phe Asn Lys Ala Met Ile Asn His Thr Thr Val Val Met Lys Lys Ile
            180                 185                 190

ctt gaa aat tac aaa ggt ttt gag aac ctt aaa act ttg gtt gat gtt        624
Leu Glu Asn Tyr Lys Gly Phe Glu Asn Leu Lys Thr Leu Val Asp Val
        195                 200                 205

gga ggt ggt ctt gga gtt aac ctc aag atg att aca tct aaa tac ccc        672
Gly Gly Gly Leu Gly Val Asn Leu Lys Met Ile Thr Ser Lys Tyr Pro
    210                 215                 220

aca att aag ggc act aat ttt gat ttg cca cat gtt gtt caa cat gcc        720
Thr Ile Lys Gly Thr Asn Phe Asp Leu Pro His Val Val Gln His Ala
225                 230                 235                 240

cct tcc tat cct ggg gtg gaa cat gtt ggg gga gat atg ttt gaa agt        768
Pro Ser Tyr Pro Gly Val Glu His Val Gly Gly Asp Met Phe Glu Ser
                245                 250                 255

gtt cca gaa gga gat gct att ttt atg aag tgg att ctt cat gac tgg        816
Val Pro Glu Gly Asp Ala Ile Phe Met Lys Trp Ile Leu His Asp Trp
        260                 265                 270

agt gat agt cac aac ctc aag ttg cta aag aac tgc tac aag gct cta        864
Ser Asp Ser His Asn Leu Lys Leu Leu Lys Asn Cys Tyr Lys Ala Leu
        275                 280                 285

cca gac aat gga aag gtg att gtt gtt gag gcc att tta cca gtg aaa        912
Pro Asp Asn Gly Lys Val Ile Val Val Glu Ala Ile Leu Pro Val Lys
    290                 295                 300

cca gac att gac acc gca gtg gtt ggc gtt tcg caa tgt gat ttg atc        960
Pro Asp Ile Asp Thr Ala Val Val Gly Val Ser Gln Cys Asp Leu Ile
305                 310                 315                 320

atg atg gct caa aat cct gga ggc aaa gag cga tcg gaa gag gag ttt       1008
Met Met Ala Gln Asn Pro Gly Gly Lys Glu Arg Ser Glu Glu Glu Phe
                325                 330                 335
```

```
cga gcc ttg gct act gaa gct gga ttc aaa ggc gtt aac tta ata tgt    1056
Arg Ala Leu Ala Thr Glu Ala Gly Phe Lys Gly Val Asn Leu Ile Cys
        340                 345                 350

tgt gtc tgt aat ttt tgg gtc atg gaa ttc tgc aag tag              1095
Cys Val Cys Asn Phe Trp Val Met Glu Phe Cys Lys  *
        355                 360
```

<210> 4
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide No. 1

<400> 4
cgtttcgcaa tgtgatttga tc   22

<210> 5
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide No. 2

<400> 5
ctcaaaatga catcctttca tac   23

<210> 6
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide No. 3

<400> 6
ctgaagatgt caatagttgc atggc   25

<210> 7
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide No. PAS1

<400> 7
ggtctagagg gccttttaga gtgttttgt tag   33

<210> 8
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide No. PS1

<400> 8
aaagtcgacc gtccacctgt gccaacaat   29

<210> 9
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide No. PS2

<400> 9
tgtttggtgt tatgcttccg tcct   24

<210> 10
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide No. PS3

<400> 10
aaaaagcttt tttaggatgg agtacagcc   29

<210> 11
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide No. PS4

<400> 11
tttaagctta aagagaacca gacaatatt   29

<210> 12
<211> 354
<212> DNA
<213> Phytophthora megasperma

<220>
<221> CDS
<222> (1)...(354)

<221> CDS
<222> (1)...(60)
<223> Preprotein

<400> 12
atg aac ttc acc gct ctg ctc gct gcc gtc gcc gcc gcc ttg gtc gga   48
Met Asn Phe Thr Ala Leu Leu Ala Ala Val Ala Ala Ala Leu Val Gly
 1   5   10   15

```
tct gcc aac gcc acc gcg tgc acc gcc acc cag cag acc gct gcg tac      96
Ser Ala Asn Ala Thr Ala Cys Thr Ala Thr Gln Gln Thr Ala Ala Tyr
            20                      25                      30

aag aca ctc gtg agc atc ctg tcg gac gcg tcg ttc aac aag tgc tct     144
Lys Thr Leu Val Ser Ile Leu Ser Asp Ala Ser Phe Asn Lys Cys Ser
            35                      40                      45

acg gat tcg ggc tac tcc atg ctg acg gcc aag gcc ctc ccc acc acg     192
Thr Asp Ser Gly Tyr Ser Met Leu Thr Ala Lys Ala Leu Pro Thr Thr
            50                      55                      60

gcg cag tac aag ctc atg tgc gcg tcc acg gca tgc aac acc atg atc     240
Ala Gln Tyr Lys Leu Met Cys Ala Ser Thr Ala Cys Asn Thr Met Ile
65                      70                      75                  80

aag aag atc gtg acg ctg aac ccg ccc aac tgc gac ctg acg gtg ccc     288
Lys Lys Ile Val Thr Leu Asn Pro Pro Asn Cys Asp Leu Thr Val Pro
                    85                      90                      95

acg agc ggc ctg gtg ctc aac gtg tac tcg tac gcg aac ggc ttc tcg     336
Thr Ser Gly Leu Val Leu Asn Val Tyr Ser Tyr Ala Asn Gly Phe Ser
                100                     105                     110

gac aag tgc tcg tcg ctg                                             354
Asp Lys Cys Ser Ser Leu
            115


<210> 13
<211> 294
<212> DNA
<213> Phytophthora megasperma

<220>
<221> CDS
<222> (1)...(294)

<400> 13
acc gcg tgc acc gcc acc cag cag acc gct gcg tac aag aca ctc gtg      48
Thr Ala Cys Thr Ala Thr Gln Gln Thr Ala Ala Tyr Lys Thr Leu Val
1           5                       10                      15

agc atc ctg tcg gac gcg tcg ttc aac aag tgc tct acg gat tcg ggc      96
Ser Ile Leu Ser Asp Ala Ser Phe Asn Lys Cys Ser Thr Asp Ser Gly
            20                      25                      30

tac tcc atg ctg acg gcc aag gcc ctc ccc acc acg gcg cag tac aag     144
Tyr Ser Met Leu Thr Ala Lys Ala Leu Pro Thr Thr Ala Gln Tyr Lys
            35                      40                      45

ctc atg tgc gcg tcc acg gca tgc aac acc atg atc aag aag atc gtg     192
Leu Met Cys Ala Ser Thr Ala Cys Asn Thr Met Ile Lys Lys Ile Val
    50                      55                      60

acg ctg aac ccg ccc aac tgc gac ctg acg gtg ccc acg agc ggc ctg     240
```

```
Thr Leu Asn Pro Pro Asn Cys Asp Leu Thr Val Pro Thr Ser Gly Leu
   65                  70                  75                  80

gtg ctc aac gtg tac tcg tac gcg aac ggc ttc tcg gac aag tgc tcg    288
Val Leu Asn Val Tyr Ser Tyr Ala Asn Gly Phe Ser Asp Lys Cys Ser
                 85                  90                  95

tcg ctg                                                             294
Ser Leu
```

```
<210> 14
<211> 1620
<212> DNA
<213> Artificial Sequence

<220>
<221> promoter
<222> (1)...(1263)
<223> COMTII promoter

<221> CDS
<222> (1264)...(1620)
<223> CDS megaspermine

<223> Synthetic construct

<400> 14
cgtccacctg tgccaacaat atagagacaa tttgctcgta tagtcagaaa gagtgttta  60
cttttttagtt gcttttttagt gaatctactc ggtataaagt taaattagtg ggtcaataag 120
tcgggtgaat agttaaagaa aacagtggtg agtttagctg tcaaataatt tcttcttttt 180
cttgttttca cattagaaat caaaataaaa cacaagcttt ttgtatttat tttaacacaa 240
gctaattata tgtttatatg ctggttaggt gaagtaaagc atgttatatg aggaaagtac 300
gaagaaaatg tgccaattgt cgtgtacagc aaagcagcca gcacaagcaa attcgcactt 360
gataagtggc taagtccact ttctagtgga cctagtggtt cactaacttt taccaaaaag 420
gcaataattt gcaattcaaa aagaaaaaag gaaaaaagaa aactagacag actttaacac 480
accaactccc acaggaagca acaatgcaac tcacaaaagg aaaccgagtt tttccgcgac 540
ggatctagaa tttgggttca ttctttacgc tttttcgtat taaactcatt atatttgtat 600
aattatgggt ttatatttttt tatttattgt aattttttgta aaattttata tataagtgta 660
tactccacgt ctccggatac tacattagcc tctagggttc ttaatactct tgttaaattg 720
tccaggctcc aaacgcatgt tcgtttcaat tttaacggat gtttccgaac aactccaaat.780
gttcaatgtt aggtgtgttt ggtgttaagc ttccgtccta ggttaataga atagataatt 840
gttgtttctt atatagtttt gaacaatcgt cgccataaac taattttttag gatggaagct 900
aattttttagg atggagtaca gcctaaggtt aaaatataac tataaaaaat atccataaaa 960
ggtgaaattt aattagtaac atgaaaagat aaaactagtg ttatcggtca aactttcaaa 1020
agagaaagaa ataactagac aaacttcaac aaccaacctg cccaacatgc tactgtgcaa 1080
ttgaaaaata aacaaaagag aaccagacaa tatttcaacc aatattccat caagaaaacc 1140
aattatgaca attcttaacc aaagtcacaa ctaacactta taaaaagcac taactcaact 1200
gtacatgatt gtgaagccta acaaaaacac tctaaaaggc ctctagagga tcccccggggt 1260
acc atg aac ttc acc gct ctg ctc gct gcc gtc gcc gcc gcc ttg gtc    1308
    Met Asn Phe Thr Ala Leu Leu Ala Ala Val Ala Ala Ala Leu Val
     1               5                  10                  15

gga tct gcc aac gcc acc gcg tgc acc gcc acc cag caa acc gct gcg    1356
Gly Ser Ala Asn Ala Thr Ala Cys Thr Ala Thr Gln Gln Thr Ala Ala
                 20                  25                  30
```

```
tac aaa aca ctc gtg agc atc ctg tcg gac gcg tcg ttc aac aag tgc    1404
Tyr Lys Thr Leu Val Ser Ile Leu Ser Asp Ala Ser Phe Asn Lys Cys
            35                  40                  45

tct acg gat tcg ggc tac tcc atg ctg acg gcc aag gcc ctc ccc acc    1452
Ser Thr Asp Ser Gly Tyr Ser Met Leu Thr Ala Lys Ala Leu Pro Thr
            50                  55                  60

acg gcg cag tac aag ctc atg tgc gcg tcc acg gca tgc aac acc atg    1500
Thr Ala Gln Tyr Lys Leu Met Cys Ala Ser Thr Ala Cys Asn Thr Met
        65                  70                  75

atc aaa aaa atc gtg acg ctg aac ccg ccc aac tgc aac ctg acg gtg    1548
Ile Lys Lys Ile Val Thr Leu Asn Pro Pro Asn Cys Asn Leu Thr Val
 80                  85                  90                  95

ccc acg agc ggc ctg gtg ctc aac gtg tac tcg tac cca aac ggc ttc    1596
Pro Thr Ser Gly Leu Val Leu Asn Val Tyr Ser Tyr Pro Asn Gly Phe
                100                 105                 110

tcg gac aag tgc tcg tcg ctg taa                                    1620
Ser Asp Lys Cys Ser Ser Leu  *
                115


<210> 15
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide PAS2

<400> 15
cgcggatccc cttttagagt gtttttgtta ggc                               33

<210> 16
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide PS5

<400> 16
acgcgtcgac gttagggaca atctatagtg tcac                              34

<210> 17
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide PS6

<400> 17
```

```
acgcgtcgac gctccgagga tttggctgtc gcgg                              34

<210> 18
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide PS7

<400> 18
acgcgtcgac gctggttagg tgaagtaaag catg                              34

<210> 19
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide PS8

<400> 19
acgcgtcgac gcatgttata tgaggaaagt acg                               33

<210> 20
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide PS9

<400> 20
acgcgtcgac gcagccagca caagcaaatt cgc                               33

<210> 21
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide PS10

<400> 21
acgcgtcgac gactttaaca caccaactcc c                                 31

<210> 22
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide PS11

<400> 22
acgcgtcgac cggatctaga atttgggttc attc                              34
```

<210> 23
<211> 35
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide PS12

<400> 23
acgcgtcgac gtgtatactc cacgtctccg gatac   35

<210> 24
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide PS13

<400> 24
acgcgtcgac gttcaatgtt aggtgtgttt gg   32

<210> 25
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide PAS3

<400> 25
cgcggatccg cttaacacca aacacaccta acattg   36

<210> 26
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide PS14

<400> 26
acgcgtcgac cagtggtgag tttagctgtc   30

SEQUENCE LISTING

(1) GENERAL INFORMATION:

    (i)   APPLICANT:
        (A)   NAME: RHOBIO
        (B)   STREET: 14-20 Rue Pierre BAIZET
        (C)   CITY: LYONS
        (E)   COUNTRY: France
        (F)   POSTAL CODE: 69009

    (ii) TITLE OF INVENTION: Inducible COMT II
promoter, chimeric gene comprising it and transformed
plants

    (iii) NUMBER OF SEQUENCES: 26

    (iv) COMPUTER READABLE FORM:
        (A)   MEDIUM TYPE: Floppy disk
        (B)   COMPUTER: IBM PC compatible
        (C)   OPERATING SYSTEM: PC-DOS/MS-DOS
        (D)   SOFTWARE: PatentIn Release #1.0,   #1.30
            (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

    (i)   SEQUENCE CHARACTERISTICS:
        (A)   LENGTH: 1863 base pairs
        (B)   TYPE: nucleotide
        (C)   STRANDEDNESS: double
        (D)   TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (ix) FEATURE:
        (A)   NAME/KEY: misc_signal
        (B)   LOCATION: 667..672
        (D)   OTHER INFORMATION:/function = "inverted
            W box"

    (ix) FEATURE:
        (A)   NAME/KEY: misc_signal
        (B)   LOCATION: 820..830
        (D)   OTHER INFORMATION:/function = "inverted
            L box"

    (ix) FEATURE:
        (A)   NAME/KEY: enhancer
        (B)   LOCATION: 845..852

    (ix) FEATURE:

        (A)   NAME/KEY: misc_signal
        (B)   LOCATION: 1034..1047
        (D)   OTHER INFORMATION:/function = "P box"

(ix) FEATURE:
        (A)   NAME/KEY: misc_signal
        (B)   LOCATION: 1221..1226
        (D)   OTHER INFORMATION:/function = "G box"

(ix) FEATURE:
        (A)   NAME/KEY: misc_signal
        (B)   LOCATION: 1343..1356
        (D)   OTHER INFORMATION:/function = "inverted
               L box"

(ix) FEATURE:
        (A)   NAME/KEY: misc_signal
        (B)   LOCATION: 1369..1374
        (D)   OTHER INFORMATION:/function = "A box"

(ix) FEATURE:
        (A)   NAME/KEY: misc_signal
        (B)   LOCATION: 1377..1382
        (D)   OTHER INFORMATION:/function = "GT box"

(ix) FEATURE:
        (A)   NAME/KEY: misc_signal
        (B)   LOCATION: 1483..1488
        (D)   OTHER INFORMATION:/function = "GT box"

(ix) FEATURE:
        (A)   NAME/KEY: misc_signal
        (B)   LOCATION: 1562..1567
        (D)   OTHER INFORMATION:/function = "inverted
               W box"

(ix) FEATURE:
        (A)   NAME/KEY: misc_signal
        (B)   LOCATION: 1600..1614
        (D)   OTHER INFORMATION:/function = "L box"

(ix) FEATURE:
        (A)   NAME/KEY: CAAT_signal
        (B)   LOCATION: 1675..1679

(ix) FEATURE:
        (A)   NAME/KEY: misc_signal
        (B)   LOCATION: 1681..1690
        (D)   OTHER INFORMATION:/function = "E box"

(ix) FEATURE:

        (A)   NAME/KEY: CAAT_signal
        (B)   LOCATION: 1695..1699

   (ix) FEATURE:
        (A)   NAME/KEY: TATA_signal
        (B)   LOCATION: 1735..1739

   (ix) FEATURE:
        (A)   NAME/KEY: transcription origin
        (B)   LOCATION: 1772

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
AAAGTTAGGG ACAATCTATA GTGTCACAAA GTTGCTTATG GCTTTTGGTT CAGATAAAGA   60

AAAAGAACAG CATTTTAATT TGTGAAGATT AGTCTGAGCA GAATTTCATT GTATCTAGAA   120

AGAAATTGAA AAAAGAAATA TTCTATTTCA CTATTATGTT AGGTGCAACT ATATCATCAC   180

CATGGAAAAG CCGGAGTAAA AAGAGAACGT AGAGGAGATT TCATGATTTG ATTGAGAATA   240

TAATATATTA TTTTTTTGTA ATTCCACACA AAGATTAAGA AAATGATCTG ATCAATGATG   300

GCTCCGAGGA TTTGGCTGTC GCGGGAACTA TGACATTAAT ATAAATTTGT CGCTGCCTAT   360

AAAGACCCTA TCTATCTATC TATCTATCTA TATATATATA TATATATATA TATATATATA   420

TATATATATA TATATATATA TATATATATA TATATATAAG CGCTAATATT TGATTATTTT   480

TTAAAAATAT TTATAAGTAT ATATGAAATT TTTGACGAAA TTTTTGTGTG ACCGTGACCC   540

CTCAACCTAT AGTGTGCGTC CACCTGTGCC AACAATATAG AGACAATTTG CTCGTATAGT   600

```
CAGAAAGAGT GTTTTACTTT TTAGTTGCTT TTTAGTGAAT CTACTCGGTA TAAAGTTAAA      660

TTAGTGGGTC AATAAGTCGG GTGAATAGTT AAAGAAAACA GTGGTGAGTT TAGCTGTCAA      720

ATAATTTCTT CTTTTTCTTG TTTTCACATT AGAAATCAAA ATAAAACACA AGCTTTTTGT      780

ATTTATTTTA ACACAAGCTA ATTATATGTT TATATGCTGG TTAGGTGAAG TAAAGCATGT      840

TATATGAGGA AAGTACGAAG AAAATGTGCC AATTGTCGTG TACAGCAAAG CAGCCAGCAC      900

AAGCAAATTC GCACTTGATA AGTGGCTAAG TCCACTTTCT AGTGGACCTA GTGGTTCACT      960

AACTTTTACC AAAAAGGCAA TAATTTGCAA TTCAAAAAGA AAAAAGGAAA AAAGAAAACT     1020

AGACAGACTT TAACACACCA ACTCCCACAG GAAGCAACAA TGCAACTCAC AAAAGGAAAC     1080

CGAGTTTTTC CGCGACGGAT CTAGAATTTG GGTTCATTCT TTACGCTTTT TCGTATTAAA     1140

CTCATTATAT TTGTATAATT ATGGGTTTAT ATTTTTTATT TATTGTAATT TTTGTAAAAT     1200

TTTATATATA AGTGTATACT CCACGTCTCC GGATACTACA TTAGCCTCTA GGGTTCTTAA     1260

TACTCTTGTT AAATTGTCCA GGCTCCAAAC GCATGTTCGT TTCAATTTTA ACGGATGTTT     1320

CCGAACAACT CCAAATGTTC AATGTTAGGT GTGTTGGTG TTAAGCTTCC GTCCTAGGTT      1380

AATAGAATAG ATAATTGTTG TTTCTTATAT AGTTTGAAC AATCGTCGCC ATAAACTAAT      1440

TTTTAGGATG GAAGCTAATT TTTAGGATGG AGTACAGCCT AAGGTTAAAA TATAACTATA     1500

AAAAATATCC ATAAAAGGTG AAATTTAATT AGTAACATGA AAAGATAAAA CTAGTGTTAT     1560

CGGTCAAACT TTCAAAGAG AAAGAAATAA CTAGACAAAC TTCAACAACC AACCTGCCCA      1620

ACATGCTACT GTGCAATTGA AAAATAAACA AAAGAGAACC AGACAATATT TCAACCAATA     1680

TTCCATCAAG AAAACCAATT ATGACAATTC TTAACCAAAG TCACAACTAA CACTTATAAA     1740

AAGCACTAAC TCAACTGTAC ATGATTGTGA AGCCTAACAA AAACACTCTA AAAGGAAAAG     1800

ACTACGAGAA TAATTACACT ACAACTCTTA TAGCTAATTC TTGTCTCAAG ATTTTCAGCT     1860

ATG                                                                  1863
```

(2)   INFORMATION FOR SEQ ID NO: 2:

    (i)   SEQUENCE CHARACTERISTICS:
       (A)   LENGTH: 5371 base pairs
       (B)   TYPE: nucleotide
       (C)   STRANDEDNESS: double
       (D)   TOPOLOGY: linear

    (ii)   MOLECULE TYPE: DNA (genomic)

    (ix)   FEATURE:
       (A)   NAME/KEY: promoter
       (B)   LOCATION: 1..1860

    (ix)   FEATURE:
       (A)   NAME/KEY: transcription origin
       (B)   LOCATION: 1772

    (ix)   FEATURE:
       (A)   NAME/KEY: exon
       (B)   LOCATION: 1861..2281

        (ix) FEATURE:
              (A)   NAME/KEY: intron
              (B)   LOCATION: 2282..3633

        (ix) FEATURE:
              (A)   NAME/KEY: exon
              (B)   LOCATION: 3634..3944

        (ix) FEATURE:
              (A)   NAME/KEY: intron
              (B)   LOCATION: 3945..4726

        (ix) FEATURE:
              (A)   NAME/KEY: exon
              (B)   LOCATION: 4727..5089

        (ix) FEATURE:
              (A)   NAME/KEY: terminator
              (B)   LOCATION: 5090..5371

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

```
AAAGTTAGGG ACAATCTATA GTGTCACAAA GTTGCTTATG GCTTTTGGTT CAGATAAAGA        60

AAAAGAACAG CATTTTAATT TGTGAAGATT AGTCTGAGCA GAATTTCATT GTATCTAGAA       120

AGAAATTGAA AAAAGAAATA TTCTATTTCA CTATTATGTT AGGTGCAACT ATATCATCAC       180

CATGGAAAAG CCGGAGTAAA AAGAGAACGT AGAGGAGATT TCATGATTTG ATTGAGAATA       240

TAATATATTA TTTTTTTGTA ATTCCACACA AAGATTAAGA AAATGATCTG ATCAATGATG       300

GCTCCGAGGA TTTGGCTGTC GCGGGAACTA TGACATTAAT ATAAATTTGT CGCTGCCTAT       360

AAAGACCCTA TCTATCTATC TATCTATCTA TATATATATA TATATATATA TATATATATA       420

TATATATATA TATATATATA TATATATATA TATATATAAG CGCTAATATT TGATTATTTT       480

TTAAAAATAT TTATAAGTAT ATATGAAATT TTTGACGAAA TTTTTGTGTG ACCGTGACCC       540

CTCAACCTAT AGTGTGCGTC CACCTGTGCC AACAATATAG AGACAATTTG CTCGTATAGT       600

CAGAAAGAGT GTTTTACTTT TTAGTTGCTT TTTAGTGAAT CTACTCGGTA TAAAGTTAAA       660

TTAGTGGGTC AATAAGTCGG GTGAATAGTT AAAGAAAACA GTGGTGAGTT TAGCTGTCAA       720

ATAATTTCTT CTTTTTCTTG TTTTCACATT AGAAATCAAA ATAAAACACA AGCTTTTTGT       780

ATTTATTTTA ACACAAGCTA ATTATATGTT TATATGCTGG TTAGGTGAAG TAAAGCATGT       840

TATATGAGGA AAGTACGAAG AAAATGTGCC AATTGTCGTG TACAGCAAAG CAGCCAGCAC       900

AAGCAAATTC GCACTTGATA AGTGGCTAAG TCCACTTTCT AGTGGACCTA GTGGTTCACT       960

AACTTTTACC AAAAAGGCAA TAATTTGCAA TTCAAAAAGA AAAAAGGAAA AAAGAAAACT      1020

AGACAGACTT TAACACACCA ACTCCCACAG GAAGCAACAA TGCAACTCAC AAAAGGAAAC      1080

CGAGTTTTTC CGCGACGGAT CTAGAATTTG GGTTCATTCT TTACGCTTTT TCGTATTAAA      1140

CTCATTATAT TTGTATAATT ATGGGTTTAT ATTTTTTATT TATTGTAATT TTTGTAAAAT      1200

TTTATATATA AGTGTATACT CCACGTCTCC GGATACTACA TTAGCCTCTA GGGTTCTTAA      1260
```

```
TACTCTTGTT AAATTGTCCA GGCTCCAAAC GCATGTTCGT TTCAATTTTA ACGGATGTTT      1320

CCGAACAACT CCAAATGTTC AATGTTAGGT GTGTTTGGTG TTAAGCTTCC GTCCTAGGTT      1380

AATAGAATAG ATAATTGTTG TTTCTTATAT AGTTTTGAAC AATCGTCGCC ATAAACTAAT      1440

TTTTAGGATG GAAGCTAATT TTTAGGATGG AGTACAGCCT AAGGTTAAAA TATAACTATA      1500

AAAAATATCC ATAAAAGGTG AAATTTAATT AGTAACATGA AAAGATAAAA CTAGTGTTAT      1560

CGGTCAAACT TTCAAAAGAG AAAGAAATAA CTAGACAAAC TTCAACAACC AACCTGCCCA      1620

ACATGCTACT GTGCAATTGA AAAATAAACA AAAGAGAACC AGACAATATT TCAACCAATA      1680

TTCCATCAAG AAAACCAATT ATGACAATTC TTAACCAAAG TCACAACTAA CACTTATAAA      1740

AAGCACTAAC TCAACTGTAC ATGATTGTGA AGCCTAACAA AAACACTCTA AAAGGAAAAG      1800

ACTACGAGAA TAATTACACT ACAACTCTTA TAGCTAATTC TTGTCTCAAG ATTTTCAGCT      1860

ATGGAATCCT CAACCAAAAG CCAAATACCA ACACAATCAG AAGAAGAGCG TAACTGCACA      1920

TATGCCATGC AACTATTGTC ATCTTCAGTC CTCCCCTTTG TGTTGCATTC AACAATTCAA      1980

TTGGAAGTTT TTGAGATATT AGCCAAATCT AATGACACTA AACTTTCTGC TTCTCAAATT      2040

GTTTCTCAAA TTCCTAACTG CACAAAACCT GAAGCACCTA CTATGTTAAA TAGGATGCTT      2100

TATGTCTTGG CTAGTTACTC CTTGTTTACT TGTTCCATTG TTGAAGATGA AAAAAATAAT      2160

GGGGGCCAAA AAAGAGTGTA TGGTTTGTCA CAAGTGGGAA AATTCTTTGT TAAAAATGAA      2220

AATGGTGCAT CAATGGGGCC ACTTTTGGCT TTGCTTCAAA ATAAAGTATT CATAAACAGC      2280

TGGTAAGTTT TGTCCTACTG TGTATTCTTT TTGCAGTGGC TGTATTGATT GGTTGCCTTT      2340

TTCACAAGAC AAGATTCTTA AGTTTTATTA CTTGTCGATT TATGTTAGTC GTATGTGCTA      2400

GTGTTATTAT TCTCCATCTG ATCCTTTTAT TGGTCACTTT ACCTAAAAAT ATTGTTACAA      2460

AACATTTGTC CTTCTAGAAA ATCAGGTATT ATTAATTTTT CAATTCCATC TTTATTACTC      2520

CAATAGTGAA TATGGTTATT AATTAGTGTT TTAAGGAAGA TGTAAGGATA ATTTAATCAA      2580

ATAGGATTTA TTATTAATGT TGTCAAAGAT TCTGGTGGAT GGATCGGAGA AAATTTCTTC      2640

ATCTTAATCA GAGTTTGATG TTCGAGCCAC AGGAATGAAT TTGTTTTTAA TAGGGAGTAT      2700

TTTCTCTTTG AATAGACCTT ACACAATAAA AGGACAACCC GGTACACTAA GCTTCCGTTA      2760

TGCGCGGGGT TCGGGGAAAG GACCGCATCA CCAGGTCTAT TGTACGCAGC GTTACCCAAC      2820

GTGAATCTAA ATTAATGAGA CTAAAAAATG GAACCCAACA CCAGTGAAAA CCAAAAAAAG      2880

AAGCAAACTT TAGTGGATGG CTTGGAAAGA TCTTTCTTCT TGAATAACTT GGAGCGCTAT      2940

ATATTAAGGC GTCGCAGCCG TTAGATACTT TCAAGAAGAA AGCTAAAAAA TGTTTTAAAG      3000

·TTACGGCGCT AGAATAATGA AATTTCTCTA TATATATAAT TCAAAAGTTA ATAATTTATT      3060

CTCTTAACTT AAATCTATAT TATAAAACTA TATTAAGTAA CTTCTGCCTA ATTTATAATA      3120

TACAACTAAT GTTTTGAGAA AACAAAATAA CAACAACATC AAACCCAATG AAATCCCACA      3180

AGTAGAGTTT GGGGAGGATA GTGTGTACGG AGACCTTACC CCTACCTTAT AAAGTTAAAG      3240
```

```
AGGCTGTTTT CGAAAGACTC TCGGCTCAAG AACATTAAAA ATTTGAGAAA ACAAAATATA    3300

AATTCAAAAC CTATATTAAG TTTATAATCC ATGGTATATT ATATTGGCTT AGTAATCTGA    3360

AATGAAAGAT TTATGTTTGA CTCCTCTAAA CTTGTTTTTA ATGCAAAAGA GGCACAACAT    3420

ATATATTATA AGTATCTTTT TTTGGTTTCC CACTGTGGCC GCTAAATTCG GATTCGCTGG    3480

AAGTGTCACA TTGTTGGAGA TGGGGGCAAC GCTCACAACA AAGACGATTC TATAATTAGT    3540

GTTCGAACCT GAAATTTTAG TTAAAGATAA AGAAGTACTT ACCATAATGG TAGATATGAT    3600

CATATCTGAC TCTCTTTCTA ATTTCAAATT ACAGGTTTGA ACTAAAAGAT GCAGTTCTTG    3660

AAGGAGGAGT TCCATTTGAC AGGGTACACG GTGTGCATGC ATTTGAATAT CCAAAATCGG    3720

ACCCAAAATT CAATGATGTT TTCAACAAGG CAATGATCAA TCACACAACT GTAGTCATGA    3780

AAAAAATACT TGAAAATTAC AAAGGTTTTG AGAACCTTAA AACTTTGGTT GATGTTGGAG    3840

GTGGTCTTGG AGTTAACCTC AAGATGATTA CATCTAAATA CCCCACAATT AAGGGCACTA    3900

ATTTTGATTT GCCACATGTT GTTCAACATG CCCCTTCCTA TCCTGGTACC TTCTCTCGTT    3960

CTTATTTTGT TGTTTATTAT ATTTACTTCG ATCATCAGGT CTAGGTCTGT CAAGTTAAAT    4020

TCGTTCTCAA AAAAGTTTAT AAAGGTTTTG AACTCCATCA CCTATTGCTT TAGGATTTTG    4080

AGTTGTATGC TCTGAGTCTT GCGCATGGTA TCATAGTCAA TTTATTTAAG CTCGTTATTG    4140

CACTTGTGAA TTCTATTATA TAAGGAGTAA GCCTACCAAA AAGGAGCGAA AATATTTTCC    4200

AAAACTCTTT TTAAACCTTC CTCACCCCAT TCCCCTCTCC CCTCTCCCCC AACACCACCC    4260

ACCACCCCAA CTCCCCCGTC TTAGTTTTTT TATTTATCCT GGACTTTCTT ATATTTTATG    4320

CTTTCCTTTA ATTGAACTCT TGTAACTAAA CCATTTGCCC CCCACCCTAT AGTGTTTGCC    4380

TAAATTTTAT ATTTTTCAAA ATAATATTTT CTATTTACTA ATTAAACATT AGAAAATATT    4440

TTTCGGATTT TTTTCCACTC ACCAACCAAG CATGGGAAAA TAGTGATAAA ACTACTCATT    4500

TTTCAAAATA ATATTTTCAA GGAAAACATT TTCCTTTATA CCAAATACCC TTACTCTTGT    4560

ATACAAATCT TCATGTCGAT GATCTTGCAA TATATATACA TGTATATGTA TGATTTGATA    4620

AACCACATGA ACAAAATGGT TGAGCTCTGC GAATTGTGAT ATATGATTTG CTTATGTGTT    4680

GTGCACTATC AATTACTTAA ATTAAACTTC ATCTAATAAT ATTGCAGGGG TGGAACATGT    4740

TGGGGGAGAT ATGTTTGAAA GTGTTCCAGA AGGAGATGCT ATTTTTATGA AGTGGATTCT    4800

TCATGACTGG AGTGATAGTC ACAACCTCAA GTTGCTAAAG AACTGCTACA AGGCTCTACC    4860

AGACAATGGA AAGGTGATTG TTGTTGAGGC CATTTTACCA GTGAAACCAG ACATTGACAC    4920

CGCAGTGGTT GGCGTTTCGC AATGTGATTT GATCATGATG GCTCAAAATC CTGGAGGCAA    4980

ACAGCGATCG GAAGAGGAGT TTCGAGCCTT GGCTACTGAA GCTGGATTCA AAGGCGTTAA    5040

CTTAATATGT TGTGTCTGTA ATTTTTGGGT CATGGAATTC TGCAAGTAGA TTTCTACTGT    5100

ACATTGAGTT TCTACTACTC TTGAGTATCC ATTTATGGCA ATCTGGGACT GGAATTGCAG    5160

CTTAGTCCAG ATTGAACATT GATATTCCTA ATAATATTTC TATTATTTCC CTTGTTTATT    5220


TCTCTTGTAT GAAAGGATGT CATTTTGAGT ATTGATAATC ATGTTCTCTA GGACAGAAAT    5280
TGTAACTTTG TCCAACTTTA TTGATATTCC TAGTAAGATT TATATGACAT GTGTCTCTGG    5340
TTTGAGAAGA GTTTCAATAT CTACAGACGG G                                  5371
```

(2)   INFORMATION FOR SEQ ID NO: 3:

         (i)   SEQUENCE CHARACTERISTICS:
              (A)   LENGTH: 1095 base pairs
              (B)   TYPE: nucleotide
              (C)   STRANDEDNESS: single
              (D)   TOPOLOGY: linear

       (ii) MOLECULE TYPE: DNA

       (ix) FEATURE:
              (A)   NAME/KEY: CDS
              (B)   LOCATION: 1..1095

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

```
ATG GAA TCC TCA ACC AAA AGC CAA ATA CCA ACA CAA TCA GAA GAA GAG      48
Met Glu Ser Ser Thr Lys Ser Gln Ile Pro Thr Gln Ser Glu Glu Glu
 1               5                   10                  15

CGT AAC TGC ACA TAT GCC ATG CAA CTA TTG TCA TCT TCA GTC CTC CCC      96
Arg Asn Cys Thr Tyr Ala Met Gln Leu Leu Ser Ser Ser Val Leu Pro
                20                  25                  30

TTT GTG TTG CAT TCA ACA ATT CAA TTG GAA GTT TTT GAG ATA TTA GCC     144
Phe Val Leu His Ser Thr Ile Gln Leu Glu Val Phe Glu Ile Leu Ala
             35                  40                  45

AAA TCT AAT GAC ACT AAA CTT TCT GCT TCT CAA ATT GTT TCT CAA ATT     192
Lys Ser Asn Asp Thr Lys Leu Ser Ala Ser Gln Ile Val Ser Gln Ile
         50                  55                  60

CCT AAC TGC ACA AAA CCT GAA GCA CCT ACT ATG TTA AAT AGG ATG CTT     240
Pro Asn Cys Thr Lys Pro Glu Ala Pro Thr Met Leu Asn Arg Met Leu
 65                  70                  75                  80

TAT GTC TTG GCT AGT TAC TCC TTG TTT ACT TGT TCC ATT GTT GAA GAT     288
Tyr Val Leu Ala Ser Tyr Ser Leu Phe Thr Cys Ser Ile Val Glu Asp
                85                  90                  95

GAA AAA AAT AAT GGG GGC CAA AAA AGA GTG TAT GGT TTG TCA CAA GTG     336
Glu Lys Asn Asn Gly Gly Gln Lys Arg Val Tyr Gly Leu Ser Gln Val
                100                 105                 110

GGA AAA TTC TTT GTT AAA AAT GAA AAT GGT GCA TCA ATG GGG CCA CTT     384
Gly Lys Phe Phe Val Lys Asn Glu Asn Gly Ala Ser Met Gly Pro Leu
             115                 120                 125

TTG GCT TTG CTT CAA AAT AAA GTA TTC ATA AAC AGC TGG TTT GAA CTA     432
Leu Ala Leu Leu Gln Asn Lys Val Phe Ile Asn Ser Trp Phe Glu Leu
         130                 135                 140

AAA GAT GCA GTT CTT GAA GGA GGA GTT CCA TTT GAC AGG GTA CAC GGT     480
Lys Asp Ala Val Leu Glu Gly Gly Val Pro Phe Asp Arg Val His Gly
145                 150                 155                 160

GTG CAT GCA TTT GAA TAT CCA AAA TCG GAC CCA AAA TTC AAT GAT GTT     528
Val His Ala Phe Glu Tyr Pro Lys Ser Asp Pro Lys Phe Asn Asp Val
                165                 170                 175
```

```
TTC AAC AAG GCA ATG ATC AAT CAC ACA ACT GTA GTC ATG AAA AAA ATA          576
Phe Asn Lys Ala Met Ile Asn His Thr Thr Val Val Met Lys Lys Ile
            180                 185                 190

CTT GAA AAT TAC AAA GGT TTT GAG AAC CTT AAA ACT TTG GTT GAT GTT          624
Leu Glu Asn Tyr Lys Gly Phe Glu Asn Leu Lys Thr Leu Val Asp Val
            195                 200                 205

GGA GGT GGT CTT GGA GTT AAC CTC AAG ATG ATT ACA TCT AAA TAC CCC          672
Gly Gly Gly Leu Gly Val Asn Leu Lys Met Ile Thr Ser Lys Tyr Pro
    210                 215                 220

ACA ATT AAG GGC ACT AAT TTT GAT TTG CCA CAT GTT GTT CAA CAT GCC          720
Thr Ile Lys Gly Thr Asn Phe Asp Leu Pro His Val Val Gln His Ala
225                 230                 235                 240

CCT TCC TAT CCT GGG GTG GAA CAT GTT GGG GGA GAT ATG TTT GAA AGT          768
Pro Ser Tyr Pro Gly Val Glu His Val Gly Gly Asp Met Phe Glu Ser
                245                 250                 255

GTT CCA GAA GGA GAT GCT ATT TTT ATG AAG TGG ATT CTT CAT GAC TGG          816
Val Pro Glu Gly Asp Ala Ile Phe Met Lys Trp Ile Leu His Asp Trp
            260                 265                 270

AGT GAT AGT CAC AAC CTC AAG TTG CTA AAG AAC TGC TAC AAG GCT CTA          864
Ser Asp Ser His Asn Leu Lys Leu Leu Lys Asn Cys Tyr Lys Ala Leu
            275                 280                 285

CCA GAC AAT GGA AAG GTG ATT GTT GTT GAG GCC ATT TTA CCA GTG AAA          912
Pro Asp Asn Gly Lys Val Ile Val Val Glu Ala Ile Leu Pro Val Lys
    290                 295                 300

CCA GAC ATT GAC ACC GCA GTG GTT GGC GTT TCG CAA TGT GAT TTG ATC          960
Pro Asp Ile Asp Thr Ala Val Val Gly Val Ser Gln Cys Asp Leu Ile
305                 310                 315                 320

ATG ATG GCT CAA AAT CCT GGA GGC AAA GAG CGA TCG GAA GAG GAG TTT·        1008
Met Met Ala Gln Asn Pro Gly Gly Lys Glu Arg Ser Glu Glu Glu Phe
                325                 330                 335

CGA GCC TTG GCT ACT GAA GCT GGA TTC AAA GGC GTT AAC TTA ATA TGT        1056
Arg Ala Leu Ala Thr Glu Ala Gly Phe Lys Gly Val Asn Leu Ile Cys
            340                 345                 350

TGT GTC TGT AAT TTT TGG GTC ATG GAA TTC TGC AAG TAG                     1095
Cys Val Cys Asn Phe Trp Val Met Glu Phe Cys Lys
            355                 360
```

(2)   INFORMATION FOR SEQ ID NO: 4:

      (i)   SEQUENCE CHARACTERISTICS:
            (A)   LENGTH: 22 base pairs
            (B)   TYPE: nucleotide
            (C)   STRANDEDNESS: single
            (D)   TOPOLOGY: linear

      (ii) MOLECULE TYPE: Synthetic oligonucleotide No.1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
CGTTTCGCAA TGTGATTTGA TC   22

(2)   INFORMATION FOR SEQ ID NO: 5:

      (i)   SEQUENCE CHARACTERISTICS:
            (A)   LENGTH: 23 base pairs
            (B)   TYPE: nucleotide
            (C)   STRANDEDNESS: single

        (D)   TOPOLOGY: linear

    (ii) MOLECULE TYPE: Synthetic oligonucleotide No.2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

CTCAAAATGA CATCCTTTCA TAC   23

(2)   INFORMATION FOR SEQ ID NO: 6:

    (i)   SEQUENCE CHARACTERISTICS:
        (A)   LENGTH: 25 base pairs
        (B)   TYPE: nucleotide
        (C)   STRANDEDNESS: single
        (D)   TOPOLOGY: linear

    (ii) MOLECULE TYPE: Synthetic oligonucleotide No.3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

CTGAAGATGT CAATAGTTGC ATGGC   25

(2)   INFORMATION FOR SEQ ID NO: 7:

    (i)   SEQUENCE CHARACTERISTICS:
        (A)   LENGTH: 33 base pairs
        (B)   TYPE: nucleotide
        (C)   STRANDEDNESS: single
        (D)   TOPOLOGY: linear

    (ii) MOLECULE TYPE: Synthetic oligonucleotide PAS1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

GGTCTAGAGG GCCTTTTAGA GTGTTTTTGT TAG   33

(2)   INFORMATION FOR SEQ ID NO: 8:

    (i)   SEQUENCE CHARACTERISTICS:
        (A)   LENGTH: 29 base pairs
        (B)   TYPE: nucleotide
        (C)   STRANDEDNESS: single
        (D)   TOPOLOGY: linear

    (ii) MOLECULE TYPE: Synthetic oligonucleotide PS1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

AAAGTCGACC GTCCACCTGT GCCAACAAT   29

(2)   INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:
     (A) LENGTH: 24 base pairs
     (B) TYPE: nucleotide
     (C) STRANDEDNESS: single
     (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Synthetic oligonucleotide PS2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

TGTTTGGTGT TATGCTTCCG TCCT   24

(2) INFORMATION FOR SEQ ID NO: 10:

     (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 292 base pairs
          (B) TYPE: nucleotide
          (C) STRANDEDNESS: single
          (D) TOPOLOGY: linear

     (ii) MOLECULE TYPE: Synthetic oligonucleotide PS3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

AAAAAGCTTT TTTAGGATGG AGTACAGCC   29

(2) INFORMATION FOR SEQ ID NO: 11:

     (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 29 base pairs
          (B) TYPE: nucleotide
          (C) STRANDEDNESS: single
          (D) TOPOLOGY: linear

     (ii) MOLECULE TYPE: Synthetic oligonucleotide PS4

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

TTTAAGCTTA AAGAGAACCA GACAATATT   29

(2) INFORMATION FOR SEQ ID NO: 12:

     (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 354 base pairs
          (B) TYPE: nucleotide
          (C) STRANDEDNESS: single
          (D) TOPOLOGY: linear

     (ii) MOLECULE TYPE: DNA

(ix) FEATURE:
    (A)   NAME/KEY: CDS
    (B)   LOCATION: 1..60
    (D)   OTHER INFORMATION:/function = preprotein

(ix) FEATURE:
    (A)   NAME/KEY: CDS
    (B)   LOCATION: 61..60
    (D)   OTHER INFORMATION:/function = preprotein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

```
atg aac ttc acc gct ctg ctc gct gcc gtc gcc gcc gcc ttg gtc gga   48
Met Asn Phe Thr Ala Leu Leu Ala Ala Val Ala Ala Ala Leu Val Gly
 1           5                  10                  15

tct gcc aac gcc acc gcg tgc acc gcc acc cag cag acc gct gcg tac   96
Ser Ala Asn Ala Thr Ala Cys Thr Ala Thr Gln Gln Thr Ala Ala Tyr
        20                  25                  30

aag aca ctc gtg agc atc ctg tcg gac gcg tcg ttc aac aag tgc tct   144
Lys Thr Leu Val Ser Ile Leu Ser Asp Ala Ser Phe Asn Lys Cys Ser
        35                  40                  45

acg gat tcg ggc tac tcc atg ctg acg gcc aag gcc ctc ccc acc acg   192
Thr Asp Ser Gly Tyr Ser Met Leu Thr Ala Lys Ala Leu Pro Thr Thr
        50                  55                  60


gcg cag tac aag ctc atg tgc gcg tcc acg gca tgc aac acc atg atc   240
Ala Gln Tyr Lys Leu Met Cys Ala Ser Thr Ala Cys Asn Thr Met Ile
 65                 70                  75                  80

aag aag atc gtg acg ctg aac ccg ccc aac tgc gac ctg acg gtg ccc   288
Lys Lys Ile Val Thr Leu Asn Pro Pro Asn Cys Asp Leu Thr Val Pro
                85                  90                  95

acg agc ggc ctg gtg ctc aac gtg tac tcg tac gcg aac ggc ttc tcg   336
Thr Ser Gly Leu Val Leu Asn Val Tyr Ser Tyr Ala Asn Gly Phe Ser
             100                105                 110

gac aag tgc tcg tcg ctg                                           354
Asp Lys Cys Ser Ser Leu
        115
```

(2)   INFORMATION FOR SEQ ID NO: 13:

    (i)   SEQUENCE CHARACTERISTICS:
        (A)   LENGTH: 354 base pairs
        (B)   TYPE: nucleotide
        (C)   STRANDEDNESS: single
        (D)   TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA

    (ix) FEATURE:
        (A)   NAME/KEY: CDS
        (B)   LOCATION: 1..294

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

```
acc gcg tgc acc gcc acc cag cag acc gct gcg tac aag aca ctc gtg    48
Thr Ala Cys Thr Ala Thr Gln Gln Thr Ala Ala Tyr Lys Thr Leu Val
 1               5                   10                  15

agc atc ctg tcg gac gcg tcg ttc aac aag tgc tct acg gat tcg ggc    96
Ser Ile Leu Ser Asp Ala Ser Phe Asn Lys Cys Ser Thr Asp Ser Gly
                20                  25                  30

tac tcc atg ctg acg gcc aag gcc ctc ccc acc acg gcg cag tac aag   144
Tyr Ser Met Leu Thr Ala Lys Ala Leu Pro Thr Thr Ala Gln Tyr Lys
            35                  40                  45

ctc atg tgc gcg tcc acg gca tgc aac acc atg atc aag aag atc gtg   192
Leu Met Cys Ala Ser Thr Ala Cys Asn Thr Met Ile Lys Lys Ile Val
        50                  55                  60

acg ctg aac ccg ccc aac tgc gac ctg acg gtg ccc acg agc ggc ctg   240
Thr Leu Asn Pro Pro Asn Cys Asp Leu Thr Val Pro Thr Ser Gly Leu
65                  70                  75                  80

gtg ctc aac gtg tac tcg tac gcg aac ggc ttc tcg gac aag tgc tcg   288
Val Leu Asn Val Tyr Ser Tyr Ala Asn Gly Phe Ser Asp Lys Cys Ser
                85                  90                  95

tcg ctg                                                           294
Ser Leu
```

(2)   INFORMATION FOR SEQ ID NO: 14:

     (i)   SEQUENCE CHARACTERISTICS:
         (A)   LENGTH: 1620 base pairs
         (B)   TYPE: nucleotide
         (C)   STRANDEDNESS: single
         (D)   TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA

    (ix) FEATURE:
         (A)   NAME/KEY: COMT II promoter
         (B)   LOCATION: 1..1263

    (ix) FEATURE:
         (A)   NAME/KEY: CDS megaspermine
         (B)   LOCATION: 1264..1630

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

```
cgtccacctg tgccaacaat atagagacaa tttgctcgta tagtcagaaa gagtgtttta   60

cttttttagtt gctttttagt gaatctactc ggtataaagt taaattagtg ggtcaataag   120

tcgggtgaat agttaaagaa aacagtggtg agtttagctg tcaaataatt tcttcttttt   180

cttgttttca cattagaaat caaaataaaa cacaagcttt ttgtatttat tttaacacaa   240

gctaattata tgtttatatg ctggttaggt gaagtaaagc atgttatatg aggaaagtac   300

gaagaaaatg tgccaattgt cgtgtacagc aaagcagcca gcacaagcaa attcgcactt   360

gataagtggc taagtccact ttctagtgga cctagtggtt cactaacttt taccaaaaag   420

gcaataattt gcaattcaaa aagaaaaaag gaaaaaagaa aactagacag actttaacac   480

accaactccc acaggaagca acaatgcaac tcacaaaagg aaaccgagtt tttccgcgac   540

ggatctagaa tttgggttca ttctttacgc tttttcgtat taaactcatt atatttgtat   600

aattatgggt ttatatttt tatttattgt aattttgta aaattttata tataagtgta   660

tactccacgt ctccggatac tacattagcc tctagggttc ttaatactct tgttaaattg   720

tccaggctcc aaacgcatgt tcgtttcaat tttaacggat gtttccgaac aactccaaat   780

gttcaatgtt aggtgtgttt ggtgttaagc ttccgtccta ggttaataga atagataatt   840

gttgtttctt atatagtttt gaacaatcgt cgccataaac taatttttag gatggaagct   900

aattttttagg atggagtaca gcctaaggtt aaaatataac tataaaaaat atccataaaa   960

ggtgaaattt aattagtaac atgaaaagat aaaactagtg ttatcggtca aactttcaaa   1020

agagaaagaa ataactagac aaacttcaac aaccaacctg cccaacatgc tactgtgcaa   1080

ttgaaaaata aacaaagag aaccagacaa tatttcaacc aatattccat caagaaaacc   1140

aattatgaca attcttaacc aaagtcacaa ctaacactta taaaaagcac taactcaact   1200

gtacatgatt gtgaagccta acaaaaacac tctaaaggc ctctagagga tccccggggt   1260
```

```
acc atg aac ttc acc gct ctg ctc gct gcc gtc gcc gcc gcc ttg gtc       1308
    Met Asn Phe Thr Ala Leu Leu Ala Ala Val Ala Ala Ala Leu Val
    1               5                   10                  15

gga tct gcc aac gcc acc gcg tgc acc gcc acc cag caa acc gct gcg       1356
Gly Ser Ala Asn Ala Thr Ala Cys Thr Ala Thr Gln Gln Thr Ala Ala
            20              25                  30

tac aaa aca ctc gtg agc atc ctg tcg gac gcg tcg ttc aac aag tgc       1404
Tyr Lys Thr Leu Val Ser Ile Leu Ser Asp Ala Ser Phe Asn Lys Cys
            35                  40                  45

tct acg gat tcg ggc tac tcc atg ctg acg gcc aag gcc ctc ccc acc       1452
Ser Thr Asp Ser Gly Tyr Ser Met Leu Thr Ala Lys Ala Leu Pro Thr
            50                  55                  60

acg gcg cag tac aag ctc atg tgc gcg tcc acg gca tgc aac acc atg       1500
Thr Ala Gln Tyr Lys Leu Met Cys Ala Ser Thr Ala Cys Asn Thr Met
    65                  70                  75

atc aaa aaa atc gtg acg ctg aac ccg ccc aac tgc aac ctg acg gtg       1548
Ile Lys Lys Ile Val Thr Leu Asn Pro Pro Asn Cys Asn Leu Thr Val
 80                  85                  90                  95

ccc acg agc ggc ctg gtg ctc aac gtg tac tcg tac cca aac ggc ttc       1596
Pro Thr Ser Gly Leu Val Leu Asn Val Tyr Ser Tyr Pro Asn Gly Phe
                100                 105                 110

tcg gac aag tgc tcg tcg ctg taa                                       1620
Ser Asp Lys Cys Ser Ser Leu
            115
```

(2)   INFORMATION FOR SEQ ID NO: 15:

    (i)   SEQUENCE CHARACTERISTICS:
        (A)   LENGTH: 33 base pairs
        (B)   TYPE: nucleotide
        (C)   STRANDEDNESS: single
        (D)   TOPOLOGY: linear

    (ii) MOLECULE TYPE: Synthetic oligonucleotide PAS2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

CGCGGATCCC CTTTTAGAGT GTTTTTGTTA GGC   33

(2)   INFORMATION FOR SEQ ID NO: 16:

    (i)   SEQUENCE CHARACTERISTICS:
        (A)   LENGTH: 33 base pairs
        (B)   TYPE: nucleotide
        (C)   STRANDEDNESS: single
        (D)   TOPOLOGY: linear

    (ii) MOLECULE TYPE: Synthetic oligonucleotide PS5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

ACGCGTCGAC GTTAGGGACA ATCTATAGTG TCAC   33

(2)   INFORMATION FOR SEQ ID NO: 17:

    (i)   SEQUENCE CHARACTERISTICS:
        (A)   LENGTH: 34 base pairs
        (B)   TYPE: nucleotide
        (C)   STRANDEDNESS: single
        (D)   TOPOLOGY: linear

    (ii) MOLECULE TYPE: Synthetic oligonucleotide PS6

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

ACGCGTCGAC GCTCCGAGGA TTTGGCTGTC GCGG   34

(2)   INFORMATION FOR SEQ ID NO: 18:

    (i)   SEQUENCE CHARACTERISTICS:
        (A)   LENGTH: 34 base pairs
        (B)   TYPE: nucleotide
        (C)   STRANDEDNESS: single
        (D)   TOPOLOGY: linear

(ii) MOLECULE TYPE: Synthetic oligonucleotide PS7

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

ACGCGTCGAC GCTGGTTAGG TGAAGTAAAG CATG   34

(2)   INFORMATION FOR SEQ ID NO: 19:

        (i)   SEQUENCE CHARACTERISTICS:
               (A)   LENGTH: 33 base pairs
               (B)   TYPE: nucleotide
               (C)   STRANDEDNESS: single
               (D)   TOPOLOGY: linear

        (ii) MOLECULE TYPE: Synthetic oligonucleotide PS8

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

ACGCGTCGAC GCATGTTATA TGAGGAAAGT ACG   33

(2)   INFORMATION FOR SEQ ID NO: 20:

        (i)   SEQUENCE CHARACTERISTICS:
               (A)   LENGTH: 33 base pairs
               (B)   TYPE: nucleotide
               (C)   STRANDEDNESS: single
               (D)   TOPOLOGY: linear

        (ii) MOLECULE TYPE: Synthetic oligonucleotide PS9

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

ACGCGTCGAC GCAGCCAGCA CAAGCAAATT CGC   33

(2)   INFORMATION FOR SEQ ID NO: 21:

        (i)   SEQUENCE CHARACTERISTICS:
               (A)   LENGTH: 31 base pairs
               (B)   TYPE: nucleotide
               (C)   STRANDEDNESS: single
               (D)   TOPOLOGY: linear

        (ii) MOLECULE TYPE: Synthetic oligonucleotide PS10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

ACGCGTCGAC GACTTTAACA CACCAACTCC C   31

(2)   INFORMATION FOR SEQ ID NO: 22:

        (i)   SEQUENCE CHARACTERISTICS:
              (A)   LENGTH: 34 base pairs
              (B)   TYPE: nucleotide
              (C)   STRANDEDNESS: single
              (D)   TOPOLOGY: linear

        (ii) MOLECULE TYPE: Synthetic oligonucleotide PS11

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

ACGCGTCGAC CGGATCTAGA ATTTGGGTTC ATTC   34

(2)   INFORMATION FOR SEQ ID NO: 23:

        (i)   SEQUENCE CHARACTERISTICS:
              (A)   LENGTH: 35 base pairs
              (B)   TYPE: nucleotide
              (C)   STRANDEDNESS: single
              (D)   TOPOLOGY: linear

        (ii) MOLECULE TYPE: Synthetic oligonucleotide PS12

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

ACGCGTCGAC GTGTATACTC CACGTCTCCG GATAC   35

(2)   INFORMATION FOR SEQ ID NO: 24:

        (i)   SEQUENCE CHARACTERISTICS:
              (A)   LENGTH: 32 base pairs
              (B)   TYPE: nucleotide
              (C)   STRANDEDNESS: single
              (D)   TOPOLOGY: linear

        (ii) MOLECULE TYPE: Synthetic oligonucleotide PS13

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

ACGCGTCGAC GTTCAATGTT AGGTGTGTTT GG   32

(2)   INFORMATION FOR SEQ ID NO: 25:

        (i)   SEQUENCE CHARACTERISTICS:
              (A)   LENGTH: 36 base pairs
              (B)   TYPE: nucleotide
              (C)   STRANDEDNESS: single
              (D)   TOPOLOGY: linear

        (ii) MOLECULE TYPE: Synthetic oligonucleotide PAS3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

CGCGGATCCG CTTAACACCA AACACACCTA ACATTG   36

(2)   INFORMATION FOR SEQ ID NO: 26:

        (i)   SEQUENCE CHARACTERISTICS:
               (A)   LENGTH: 30 base pairs
               (B)   TYPE: nucleotide
               (C)   STRANDEDNESS: single
               (D)   TOPOLOGY: linear

        (ii) MOLECULE TYPE: Synthetic oligonucleotide PS14

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

ACGCGTCGAC CAGTGGTGAG TTTAGCTGTC   30